

平面パターンを用いた複数カメラシステムのキャリブレーション

植芝 俊夫

富田 文明

産業技術総合研究所

〒 305-8568 つくば市梅園 1-1-1 中央第 2

平面パターンを用いて複数のカメラを同時に校正する新しいアルゴリズムを提案する。これは Zhang による単一カメラのためのキャリブレーション法の拡張であり、2 次元位置が既知の参照点を描いたモデル平面を数カ所に置いてカメラに提示するだけで、個々のカメラの内部パラメータのみならずカメラ間の位置関係をも推定することができる。平面の動きを知る必要はないという Zhang の方法の長所が保持されるので、本手法は複数カメラシステムを校正する簡便な手段を与える。シミュレーションと実データを用いた実験により提案手法の有効性を示す。

Calibration of Multi-camera Systems Using Planar Patterns

Toshio Ueshiba

Fumiaki Tomita

National Institute of Advanced Industrial Science and Technology
AIST Tsukuba Central 2, Umezono 1-1-1, Tsukuba, 305-8568 Japan

A new calibration algorithm for multi-camera systems using planar reference patterns is proposed. The algorithm is an extension of Zhang's flexible calibration technique to the cases with multiple cameras. Rigid displacements between cameras are recovered as well as intrinsic camera parameters only by showing the cameras a model plane with known reference patterns placed at several locations. Thus the algorithm yields a simple calibration means for stereo vision systems with arbitrary number of cameras while maintaining the handiness and flexibility of Zhang's method. The validity of the proposed technique is verified by simulation and real experiments.

1 序論

画像の取得に用いるカメラの種々の特性やパラメータを求めるカメラキャリブレーションは、シーンの 3 次元復元や物体認識など、コンピュータビジョンの多くの応用で必要となる基礎的な技術である。カメラキャリブレーションには、レンズによって生じる光学的歪みを推定する光学的キャリブレーションと、焦点距離や画像中心などのカメラ固有の内部パラメータおよび

カメラの位置・姿勢を表す外部パラメータを求める幾何学的キャリブレーションがある [17] が、本稿では後者を扱い、特にステレオビジョンに代表される複数カメラシステムのキャリブレーションについて論じる。

幾何学的キャリブレーションの古典的な手法として代表的なものは、3 次元位置が既知の参照点を多数観測し、その 3 次元座標と投影像の 2 次元座標との対応関係からカメラパラメータを計算する方法である [12]。

参照点の3次元座標を正確に設定できる場合、この手法はカメラパラメータを高精度かつ安定に推定できる。また、カメラの外部パラメータは参照点位置を記述した座標系上で得られるので、カメラが複数ある場合でも、それらに同一の参照点を見せて個別にキャリブレーションを行うことにより、全カメラの位置と姿勢を共通の座標系上で求めることができる。そのため、この手法はカメラ間の相対的な位置関係が重要となる複数カメラシステムにも適用できる。しかし、参照点を3次元空間中で高精度に設定することは、特別の設備を要する手間のかかる作業である。特に、カメラの視野が大きくなると参照点を広範囲に配置せねばならず、急激にその困難さが増す。

このような問題を解決する汎用的な手法として、参照点や参照物体を一切用いないセルフキャリブレーションと呼ばれるテクニックがここ10年ほど盛んに研究されてきた。これは、未知物体上の特徴点を複数の視点から撮像し、その投影像のみからカメラの内部パラメータと外部パラメータの両方を求めるものである。研究の初期には、すべての視点において内部パラメータは不変との制約が課されていた [5, 11] ため複数カメラシステムには適用できなかったが、その後この制約をはずしてもキャリブレーションが可能であることが示され [3, 6]、応用範囲が広がった。

セルフキャリブレーションは極めて制約の少ない柔軟な手法であるが、カメラ間の相対的な位置関係にスケールの不定性が残されるため、3次元復元を行う際にも絶対的なスケールが決まらない問題がある。また、理論上解が一意に決まらない特異的なカメラ配置が存在し [8, 4]、その近傍では計算が不安定になる。さらに、利用できる制約に比べて推定すべきパラメータが多いためパラメータ間の独立性が十分でなく、特異配置から遠い場合でもしばしば数値的な安定性に欠け、特に画像中心の推定は困難であることが指摘されている [18]。

Zhang は、上述の古典的キャリブレーションとセルフキャリブレーションの中間に位置する簡便かつ安定なカメラキャリブレーション法を提案した [15, 16]。これは、参照物体として3次元的広がりを持つ立体でなく、2次元位置が既知の参照点を描いた平面(モデル平面)を用いるものである。ユーザは、モデル平面もしくはカメラを複数(2箇所以上)の異なる位置に動かして参照点を撮影し、平面と画像面との射影変換を計算してカメラパラメータを抽出する。モデル平面(またはカメラ)は自由に動かすことができ、その運動を知る必要はない。すなわち参照点の位置を2次元平面上のみで指定すれば良いので、3次元位置を与えねばならない古典的キャリブレーション法に比べてユーザ

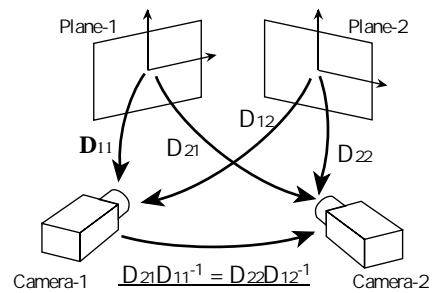


図 1: カメラ間の相対的な位置関係に関する拘束

の負担は大幅に軽減される。

Zhang の手法は極めて実用的であるが、複数カメラシステムに適用する場合に問題が生じる。カメラの位置と姿勢は、モデル平面に固定した座標系上で求められるので、複数のカメラに共通の平面を見せればこれを介してカメラ間の相対的な位置関係が計算できる。しかし、モデル平面は複数あるので、平面数だけの計算の仕方が存在する。ノイズがなければこれらはすべて一致するはずであるが、ノイズを含む実際の環境ではそうならず矛盾が生じる。

例として、2台のカメラを2枚のモデル平面でキャリブレーションを行う場合を考えよう(図1)。カメラ1のキャリブレーションを行うと、平面1, 2から見たカメラの姿勢がそれぞれ剛体変換 D_{11} , D_{12} として求まる。同様に、カメラ2についても剛体変換 D_{21} , D_{22} が計算される。すると、2つのカメラ間の相対的な位置関係は、平面1または平面2を介して計算した剛体変換 $D_{21}D_{11}^{-1}$ もしくは $D_{22}D_{12}^{-1}$ として求められるが、ノイズの存在下では両者は必ずしも一致しない。

この矛盾は、複数カメラシステムに適用する場合、Zhang の定式化ではカメラの外部パラメータが冗長であることに起因する。一般に、パラメトリゼーションが冗長であると系が満たすべき拘束が厳密に成立せず、パラメータの推定精度や計算の安定性が低下したり、推定結果を利用する段階で不整合が生じたりする。

本稿では、この問題を解決するため、Zhang のキャリブレーション法を複数カメラシステムに適用できるよう拡張する。まず、モデル平面とカメラ画像面との間の射影変換が、カメラの投影行列と平面の位置と姿勢を表す行列の積となることを示す(2節)。これにより、射影変換行列を並べた計測行列に structure-from-motion における因子分解法 [10] を適用して、カメラと平面を表す行列を取り出すことができる。その際に、射影変換行列の定数倍の不定性を補償して計測行列を分解可能にするため、平面射影変換(plane homography)の固有値に関する性質 [14] を利用する。次に、モデル

平面上で2次元ユークリッド座標系が既知であることを用いて，カメラと平面を表す行列を射影座標系からユークリッド座標系における表現に変換する．最後に，非線形最適化により全パラメータの微調整を行い，最尤推定値を求めて最終解とする（以上，3節）．シミュレーションおよび実データを用いた実験によって提案手法の有効性を検証した結果についても述べる（4節）．

2 モデル平面と画像平面の射影変換

I 台の透視投影カメラをキャリブレーションするものとして．ワールド座標系から見た i 番目 ($i = 1, \dots, I$) のカメラの位置と姿勢を，それぞれ3次元ベクトル t_i と 3×3 直交行列 R_i で表す．また，カメラの画像平面を \mathcal{I}_i とする．すると3次元座標 $X \in \mathcal{R}^3$ を持つシーンの点は， \mathcal{I}_i 上の

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \simeq \underbrace{K_i R_i^\top}_{P_i} \begin{bmatrix} I_3 & -t_i \\ & 1 \end{bmatrix} \begin{bmatrix} X \\ 1 \end{bmatrix} \quad (1)$$

なる位置に投影される¹．ただし， K_i はカメラの内部パラメータを表す 3×3 上半三角行列であり，焦点距離 k_i ，画像中心 (u_{0i}, v_{0i}) ，アスペクト比 a_i および軸の非直交性歪み s_i を用いて

$$K_i = \begin{bmatrix} a_i k_i & s_i k_i & u_{0i} \\ 0 & k_i & v_{0i} \\ 0 & 0 & 1 \end{bmatrix} \quad (2)$$

と表される [1]．カメラの投影を表す (1) 式の 3×4 行列 P_i を，カメラ行列と呼ぶ．

次に，既知の参照点を描いた J 枚のモデル平面 π_j ($j = 1, \dots, J$) を考え，これらを用いてキャリブレーションを行うものとして． π_j に固定した2次元直交座標系について，ワールド座標系から見たその横軸と縦軸の方向および原点の位置を，それぞれを3次元単位ベクトル p_j, q_j および3次元ベクトル d_j で表す (図2)． p_j, q_j の正規直交性より

$$p_j^\top p_j = q_j^\top q_j = 1, \quad p_j^\top q_j = 0 \quad (3)$$

が成り立つ．すると，2次元位置 (x, y) を持つ π_j 上の参照点は，ワールド座標系から見た場合

$$\begin{bmatrix} X \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} p_j & q_j & d_j \\ 0 & 0 & 1 \end{bmatrix}}_{Q_j} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (4)$$

なる座標に位置する．(4) 式の 4×3 行列 Q_j をモデル平面行列と呼ぶ．

¹記号 \simeq は，非零の定数倍の違いを除いて等しいことを表す．

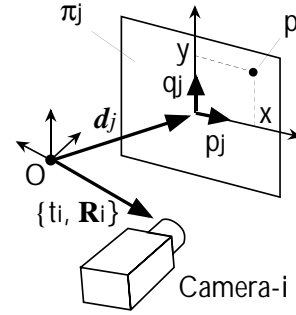


図2: カメラとモデル平面

以上 (1), (4) 式より， π_j 上の特徴点の位置 (x, y) と，その \mathcal{I}_i への投影位置 (u, v) は

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \simeq \underbrace{P_i Q_j}_{H_{ij}} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (5)$$

なる2次元射影変換で結ばれ，変換を表す 3×3 行列 H_{ij} は 3×4 カメラ行列 P_i と 4×3 モデル平面行列 Q_j の積となることがわかる．

個々の射影変換行列 H_{ij} は， π_j 上の参照点とその \mathcal{I}_i への投影との対応 $(x, y) \leftrightarrow (u, v)$ が4組以上得られれば，定数倍の不定性を除いて求めることができる．

3 キャリブレーションアルゴリズム

3.1 射影変換行列のスケーリングと因子分解

(5) 式を全てのカメラとモデル平面について連立すれば，

$$\underbrace{\begin{bmatrix} H_{11} & \cdots & H_{1J} \\ \vdots & & \vdots \\ H_{I1} & \cdots & H_{IJ} \end{bmatrix}}_W = \underbrace{\begin{bmatrix} P_1 \\ \vdots \\ P_I \end{bmatrix}}_P \underbrace{\begin{bmatrix} Q_1 & \cdots & Q_J \end{bmatrix}}_Q \quad (6)$$

を得る．左辺の $3I \times 3J$ 行列 W は計測行列と呼ばれる．(6) 式は，この行列が I 台のカメラを表す $3I \times 4$ 行列 P と J 枚のモデル平面を表す $4 \times 3J$ 行列 Q に分解でき，したがって $\text{rank } W = 4$ であることを示している．よって，すべての射影変換行列 H_{ij} がその絶対的スケールも含めて既知ならば，structure-from-motion問題における因子分解法 [10] と同様，特異値分解 (SVD) を用いて (4×4 正則行列の不定性を除いて) カメラ行列とモデル平面行列を復元することができる．

しかし，点対応から求めた射影変換行列 H_{ij} には定数倍の不定性が残るので，実際に得られるのは H_{ij} に未知定数 λ_{ij} を乗じた $\tilde{H}_{ij} = \lambda_{ij} H_{ij}$ である．一般に，この \tilde{H}_{ij} を並べて構成した計測行列 \tilde{W} はランク4に

ならず、したがって \tilde{W} をカメラ行列とモデル平面行列に分解することはできない。

同様の困難は、structure-from-motion 問題において透視投影画像に因子分解法を適用する場合 [9, 2, 13] にも生じ、2 とおりの解決策が提案されている。1 つは、計測行列ができる限りランク 4 に近くなるようなスケーリング定数 λ_{ij} を非線形最適化によって逐次的に求めることである [2, 13]。しかし、非線形最適化は計算量が多く、また真値に充分近い初期値を必要とする問題がある。そこで、我々は複数のスケーリング定数の間に成立する拘束を利用する第 2 のアプローチ [9] に類似した以下の方法を採用する。

j 番目のモデル平面 π_j によって、1 番目と i 番目の画像平面 \mathcal{I}_1 と \mathcal{I}_i の間に射影変換 $A_{ij} = H_{ij}H_{1j}^{-1}$ が定義される。そこで、2 つのモデル平面 π_1, π_j を考えることにより、 \mathcal{I}_1 からそれ自身への射影変換 $G_{ij} = A_{ij}^{-1}A_{i1}$ を定義する。この変換は、 \mathcal{I}_1 上の点を π_1 を介して \mathcal{I}_i に写像し、さらにその点を π_j を介して \mathcal{I}_1 に戻す (図 3)。最近、Zelnik-Manor らは G_{ij} は単位行列 I と 2 つの 3 次元ベクトル e, f によって $G_{ij} = I + ef^T$ という形に表され、したがって G_{ij} は 1 を 2 重固有値として持つことを示した [14]²。このことから、スケールの不定性を残した \tilde{H}_{ij} から計算した $\tilde{G}_{ij} = \tilde{H}_{1j}\tilde{H}_{ij}^{-1}\tilde{H}_{i1}\tilde{H}_{11}^{-1} = (\lambda_{i1}\lambda_{1j})/(\lambda_{11}\lambda_{ij})G_{ij}$ は、2 重固有値 $\mu_{ij} = (\lambda_{i1}\lambda_{1j})/(\lambda_{11}\lambda_{ij})$ を持つことがわかる。 μ_{ij} は \tilde{G}_{ij} から直接計算できるので、これを用いて \tilde{H}_{ij} を

$$\tilde{H}_{ij} \leftarrow \mu_{ij}\tilde{H}_{ij} \quad (i = 2, \dots, I; j = 2, \dots, J) \quad (7)$$

とスケーリングする。すると、 $\tilde{H}_{ij} = \mu_{ij}\lambda_{ij}H_{ij} = (\lambda_{i1}\lambda_{1j}/\lambda_{11})H_{ij}$ となるので、 \tilde{H}_{ij} から構成した計測行列 \tilde{W} は

$$\tilde{W} = \begin{bmatrix} \tilde{H}_{11} & \cdots & \tilde{H}_{1J} \\ \vdots & & \vdots \\ \tilde{H}_{I1} & \cdots & \tilde{H}_{IJ} \end{bmatrix} \simeq \begin{bmatrix} \lambda_{11}P_1 \\ \vdots \\ \lambda_{I1}P_I \end{bmatrix} \begin{bmatrix} \lambda_{11}Q_1 & \cdots & \lambda_{1J}Q_J \end{bmatrix} \quad (8)$$

と分解することができる。

注意 実際には、 \tilde{G}_{ij} はノイズの影響を受けるため、その固有値の 1 つ μ_{ij} が厳密に 2 重根となるとは限らな

² f に直交するベクトルが固有値 1 に対応する固有ベクトルとなることは、容易に確かめられる。 f の直交補空間は 2 次元であるから、固有値 1 の多重度は 2 となる。

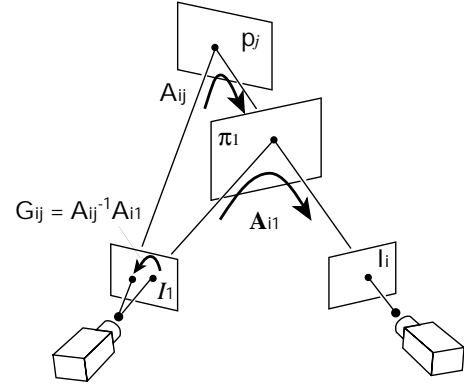


図 3: 2 台のカメラと 2 枚のモデル平面により定義される射影変換

い。ノイズの存在下で最適な μ_{ij} を計算する方法を付録 A に示した。

3.2 ユークリッド空間における表現への変換

上述の方法により、SVD を用いて計測行列 \tilde{W} が

$$\tilde{W} = \begin{bmatrix} P'_1 \\ \vdots \\ P'_I \end{bmatrix} \begin{bmatrix} Q'_1 & \cdots & Q'_J \end{bmatrix} \quad (9)$$

と分解されたとしよう。一般性を失うことなく、 P'_1 が

$$P'_1 = \begin{bmatrix} I & 0 \end{bmatrix} \quad (10)$$

なる特別な形を持つと仮定できる。(そうでなければ、 P'_1 とその零点 t'_1 から構成した 4×4 行列 $S = [P_1'^T(P_1'P_1')^{-1} \ t'_1]$ によって $P'_i \leftarrow P'_iS$ 、 $Q'_j \leftarrow S^{-1}Q'_j$ なる変換を行えば良い。)

(8), (9) 式を比べることにより、SVD によって得た P'_i と Q'_j はユークリッド空間における真のカメラ行列 P_i およびモデル平面行列 Q_j と

$$P'_i \simeq P_iT, \quad Q'_j \simeq T^{-1}Q_j \quad (11)$$

なる射影変換で結ばれることがわかる。ここで T は未知の 4×4 正則行列であり、我々はこれを決定してカメラとモデル平面の射影空間における表現をユークリッド空間における表現に変換しなければならない。

一般性を失うことなくユークリッド空間におけるワールド座標系を第 1 カメラに固定すれば、 $R_1 = I$ 、 $t_1 = 0$ となってそのカメラ行列は $P_1 = [K_1 \ 0]$ となる。これと (10), (11) 式を考慮すれば、 T は

$$T = \begin{bmatrix} K_1^{-1} & 0 \\ h^T & h \end{bmatrix} \quad (12)$$

という形になることがわかる．ここで $[\mathbf{h}^\top \ h]$ は無限遠平面を表す．

未知定数 β_j を導入して (11) の第 2 式を書き直せば

$$\mathbf{T}\mathbf{Q}'_j = \beta_j \mathbf{Q}_j \quad (13)$$

を得る． \mathbf{Q}'_j の左上 3×2 部分行列を $[\mathbf{p}'_j \ \mathbf{q}'_j]$ と書けば，(3), (12), (13) 式より，absolute conic の \mathcal{I}_1 への投影像 $\omega_1 = \mathbf{K}_1^{-\top} \mathbf{K}_1^{-1}$ に関して

$$\mathbf{p}'_j{}^\top \omega_1 \mathbf{p}'_j = \mathbf{q}'_j{}^\top \omega_1 \mathbf{q}'_j = \beta_j^2, \quad \mathbf{p}'_j{}^\top \omega_1 \mathbf{q}'_j = 0 \quad (14)$$

が成り立つ． β_j を無視すれば，上式は 3×3 対称行列 ω_1 の 6 つの独立な要素に対して 2 つの拘束を与えるので，3 枚以上 (第 1 カメラの歪みパラメータを 0 と仮定すれば 2 枚以上) のモデル平面があれば最小自乗法により ω_1 を定数倍を除いて決定することができる．これを Cholesky 分解し (3, 3) 要素を 1 に正規化することによって， \mathbf{K}_1 を求める．

次に，得られた \mathbf{K}_1 を用いて， β_j を

$$\beta_j = \sqrt{(\mathbf{p}'_j{}^\top \omega_1 \mathbf{p}'_j + \mathbf{q}'_j{}^\top \omega_1 \mathbf{q}'_j)/2} \quad (15)$$

と決定する．

最後に，(13) 式の第 4 行を取り出した

$$\begin{bmatrix} \mathbf{h}^\top & h \end{bmatrix} \mathbf{Q}'_j = \begin{bmatrix} 0 & 0 & \beta_j \end{bmatrix} \quad (16)$$

を全てのモデル平面について連立して無限遠平面 $[\mathbf{h}^\top \ h]$ を決定する．(16) 式は $[\mathbf{h}^\top \ h]$ の 4 つの要素に対して 3 つの拘束を与えるから，2 枚以上のモデル平面があれば最小自乗法によりこれを決定することができる．

以上のようにして得られた \mathbf{T} の逆行列を \mathbf{P}'_i に適用して，ユークリッド空間におけるすべてのカメラ行列 \mathbf{P}_i ($i = 1, \dots, I$) を計算する． \mathbf{P}_i の左 3×3 部分行列を QR 分解すればカメラの内部パラメータ \mathbf{K}_i と回転 \mathbf{R}_i が得られ，さらに第 4 列からカメラの位置 \mathbf{t}_i が求まる．

3.3 非線形最適化によるパラメータの微調整

これまで述べた手順により，すべてのカメラ行列 \mathbf{P}_i ($i = 1, \dots, I$) が線形演算により求められるが，その途中で数段階に渡って行われる最小自乗法による中間パラメータの計算は統計的に意味のある誤差基準に基づいたものではないので，得られた解は統計的に最適なものではない．そこで，[16] と全く同様に，非線形最適化によってパラメータの微調整を行い，最終解を決定する．

j 番目のモデル平面 π_j に N_j 個の参照点 $\mathbf{x}_{jn} = (x_{jn}, y_{jn})$ ($n = 1, \dots, N_j$) が描かれているものとしよう． \mathbf{x}_{jn} の第 i カメラの画像面 \mathcal{I}_i への像を $\mathbf{u}_{ijn} =$

(u_{ijn}, v_{ijn}) とする． \mathbf{u}_{ijn} に含まれるノイズが一様かつ等方な Gaussian ノイズであると仮定すれば，次の誤差基準を最小化することによってカメラの内部/外部パラメータの最尤推定量が求められる：

$$\sum_{i=1}^I \sum_{j=1}^J \sum_{n=1}^{N_j} \|\mathbf{u}_{ijn} - \hat{\mathbf{u}}(\mathbf{K}_i, \mathbf{R}_i, \mathbf{t}_i, \mathbf{p}_j, \mathbf{q}_j, \mathbf{d}_j; \mathbf{x}_{jn})\|^2 \rightarrow \min \quad (17)$$

ただし， $\hat{\mathbf{u}}(\mathbf{K}_i, \mathbf{R}_i, \mathbf{t}_i, \mathbf{p}_j, \mathbf{q}_j, \mathbf{d}_j; \mathbf{x}_{jn})$ は (5) 式によって \mathbf{x}_{jn} を \mathcal{I}_i へ投影する関数である．この最小化過程は非線形であるが，3.2 節までに求めた解を初期値とし，Levenberg-Marquardt 法 [7] を用いて実現できる．

現在の実装ではレンズ歪みを考慮していないが，これを $\hat{\mathbf{u}}(\cdot)$ に取り入れて推定対象とすることももちろん可能である [16, 18] ．

4 実験結果

提案したアルゴリズムの有効性を，シミュレーションおよび実データを用いた実験によって検証した．

4.1 シミュレーション

3 枚のモデル平面を用いて 3 台のカメラを校正するシミュレーション実験を行った．第 2, 3 カメラを除いて，すべてのパラメータは Zhang のレポート [15] における値と同一に設定した．カメラの内部パラメータは，3 台とも $k_i = 900$, $u_{i0} = v_{i0} = 255$, $a_i = 1.3888$, $s_i = 0.001212$ である．各モデル平面には，18mm の間隔で $10 \times 14 = 140$ 個の参照点が描かれている． i 番目のカメラの位置と姿勢はそれぞれ 3 次元ベクトル \mathbf{t}_i, θ_i で表される．ここで θ_i の長さとは向きは，それぞれ回転角と回転軸方向に等しい．同様に， j 番目のモデル平面の位置と姿勢も，2 つの 3 次元ベクトル \mathbf{d}_j, ϕ_j で表す．前節で述べたように，ワールド座標系は第 1 カメラに固定されているので， $\mathbf{t}_1 = \theta_1 = \mathbf{0}$ である．第 2, 3 カメラの位置と姿勢は，それぞれ $\{\mathbf{t}_2 = [50, 0, 0]^\top, \theta_2 = [0, -5.7296^\circ, 0]\}$, $\{\mathbf{t}_3 = [100, 0, 0]^\top, \theta_3 = [0, -11.4592^\circ, 0]\}$ と設定した．3 枚のモデル平面の位置と姿勢は，それぞれ $\{\mathbf{d}_1 = [-9, -12.5, 500]^\top, \phi_1 = [20^\circ, 0, 0]^\top\}$, $\{\mathbf{d}_2 = [-9, -12.5, 510]^\top, \phi_2 = [0, 20^\circ, 0]^\top\}$, $\{\mathbf{d}_3 = [-10.5, -12.5, 525]^\top, \phi_3 = \frac{1}{\sqrt{5}}[-30^\circ, -30^\circ, -15^\circ]^\top\}$ と設定した．

様々なレベルの等方かつ一様な Gaussian ノイズを特徴点の観測位置に加え，各レベルにつき 500 回の試行を行ってその誤差の平均値を評価した．比較のため，提案アルゴリズムと共に Zhang のアルゴリズムも実装し，全く同一のデータセットに対して実験を行った．

図 4, 5 に第 1 カメラの内部パラメータの推定誤差を示す．どのパラメータについても，提案手法は Zhang

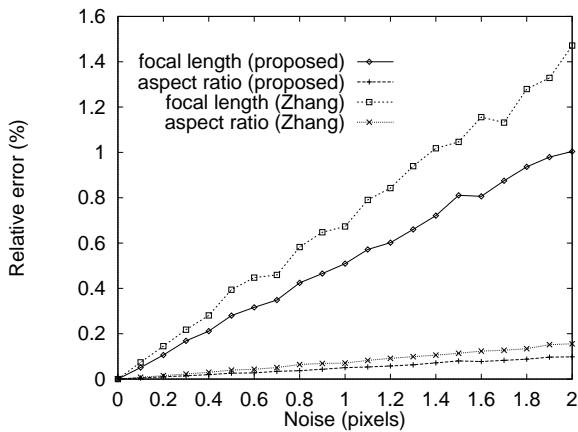


図 4: 第 1 カメラの焦点距離とアスペクト比の相対誤差

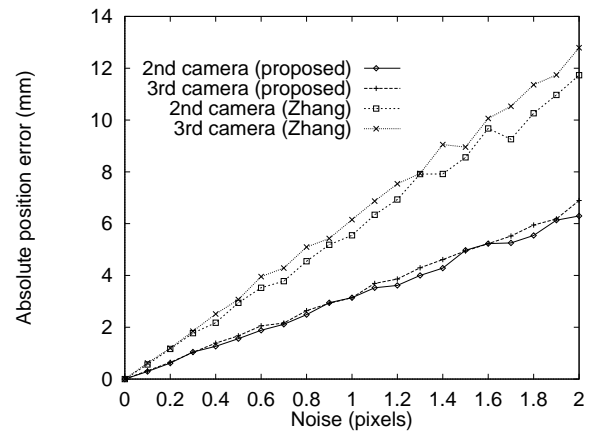


図 6: 第 2, 3 カメラの位置の絶対誤差

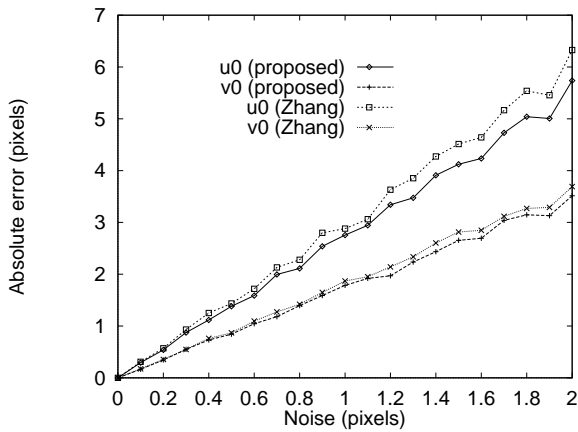


図 5: 第 1 カメラの画像中心の絶対誤差

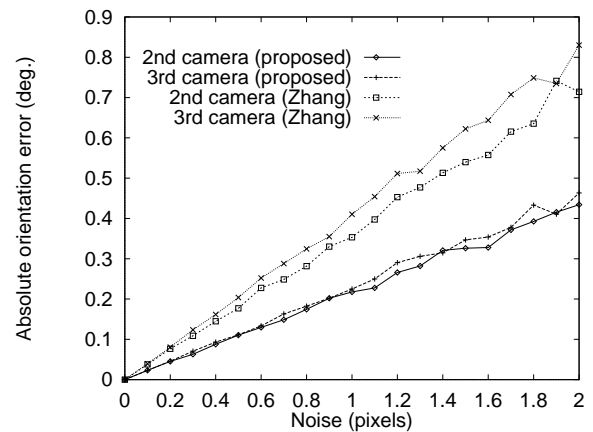


図 7: 第 2, 3 カメラの方向の絶対誤差

の方法に比べて精度が向上していることがわかる。これは、提案手法では序論で述べたカメラ間の相対的位置関係に関するパラメータの冗長性が排除されているため、ノイズへの過適合 (overfitting) が生じにくいと考えられる。

次に、カメラ間の相対的位置関係の推定誤差を、提案手法と Zhang の方法の両者について評価した。後者では、個々のカメラに対して独立にキャリブレーションを行い、1 番目のモデル平面を介してカメラ間の並進と回転を計算した。

図 6, 7 に第 1 カメラから見た第 2, 3 カメラの位置と姿勢を示す。姿勢の誤差は、真の回転行列 R_i とその推定値 \hat{R}_i に対して誤差行列 $\Delta R_i = R_i \hat{R}_i^T$ を考え、その回転角度として定義した。提案手法では、誤差が Zhang の方法による場合と比べて 50 ~ 60% に減少していることがわかる。

4.2 実データによる実験

実際のデータを用いて、焦点距離 8mm のレンズを装着した 3 台の IEEE1394 デジタルカメラを校正する実験を行った。20mm の間隔で $13 \times 9 = 117$ 個の参照点を描いたモデル平面を 6 箇所置き、それを 3 台のカメラから観測した。

図 8 は、(7) 式のスケーリングを行う前後の計測行列 \tilde{W} の特異値を表している。スケーリング後は、4 番目の特異値に比べ 5 番目の特異値が 2 桁以上小さくなっており、 \tilde{W} のランクが充分 4 に近づいていることがわかる。

表 1 に、すべてのパラメータの推定値を示す。各カメラについて、第 1 列は非線形最適化の直前の推定値を、第 2 列は最終的な推定値をそれぞれ表している。線形演算による初期値は最終解に充分近いことがわかる。

キャリブレーションされたカメラを用いて、3 眼ステレオ画像からの 3 次元復元実験を行った。ガソリンスタンドの模型を撮影した画像 (図 9(a)) にセグメント

表 1: カメラパラメータの推定値

| params | first camera | | second camera | | third camera | |
|--------------------|-----------------|-----------------|---------------------|---------------------|----------------------|----------------------|
| | initial | final | initial | final | initial | final |
| k_i | 1092.42 | 1085.84 | 1099.75 | 1093.27 | 1114.19 | 1110.85 |
| (u_{i0}, v_{i0}) | (321.06 218.16) | (318.18 217.03) | (352.36 236.03) | (354.02 235.49) | (376.67 227.68) | (373.85 225.00) |
| a_i | 1.00962 | 1.00829 | 1.00803 | 1.00809 | 1.00907 | 1.00600 |
| s_i | -0.0019157 | -0.0012213 | 0.0037976 | -0.0010745 | 0.0015064 | -0.0012511 |
| t_i (mm) | (0 0 0) | (0 0 0) | (154.12 0.89 22.66) | (154.30 0.93 22.37) | (72.36 -74.33 19.88) | (72.42 -74.11 19.62) |
| θ_i (deg.) | (0 0 0) | (0 0 0) | (-0.27 -12.55 0.55) | (-0.31 -12.40 0.57) | (-7.56 -3.73 -1.18) | (-7.56 -3.75 -1.12) |

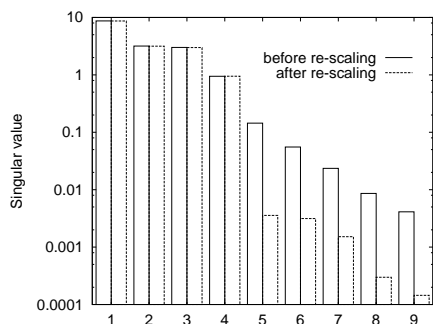
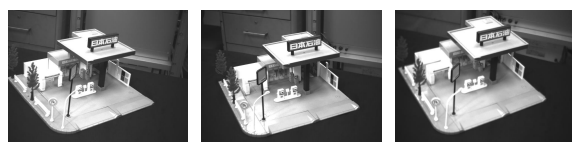
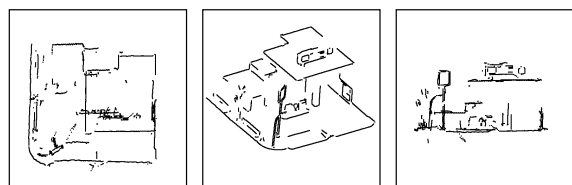


図 8: スケーリング前/後における計測行列の特異値



(a) Input image triplet



(b) Reconstructed 3D structure: top view, upper-side view and side view

図 9: 校正されたカメラにより撮像された 3 画像からの 3 次元復元結果

ベストステレオビジョンアルゴリズムを適用して得た復元結果を、異なる 3 方向から見たものを図 9(b) に示す。

5 結論

2 次元位置が既知の参照点を描いたモデル平面を用いて複数のカメラを同時に校正するキャリブレーションアルゴリズムを提案した。本手法は、Zhang による単一カメラのためのキャリブレーション法を複数カメラシステムに適用できるよう拡張したものである。本手法は、モデル平面を自由に動かすことができその運動を知る必要がないというもとのアルゴリズムの性質を引き継いでいるので、ステレオビジョンに代表され

る複数カメラシステムの簡便なキャリブレーションの手段を与える。

アルゴリズムとしての特徴は、複数のモデル平面を用いて複数のカメラを同時に校正するために structure-from-motion における因子分解法を適用したことと、その際に問題となる射影変換行列の定数倍の不定性を、2 枚の画像と 2 枚のモデル平面によって定義される平面射影変換が 2 重固有値を持つという性質を用いて解決したことの 2 点にある。シミュレーションの結果、内部/外部パラメータとも推定精度が向上し、特に Zhang の方法を複数カメラに適用した場合に生じる外部パラメータの冗長性が排除されたため、その推定誤差が約半分に軽減された。また、実データを用いた実験でも提案手法の有効性が確認された。

本手法が適用できるためには、各カメラからすべてのモデル平面が見えていなければならない。これは、カメラが広範囲に配置されていたり、シーンを取り囲むように置かれている場合に問題となる。この制約を取り除くことが今後の課題である。

A $\text{rank}(\mathbf{G} - \mu\mathbf{I}) = 1$ とする μ の求め方

$$\mathbf{G} = \mu\mathbf{I} + \mathbf{e}\mathbf{f}^T \quad (18)$$

なる形を持つ 3×3 行列 \mathbf{G} が与えられたとしよう。ここで、 μ はスカラー、 \mathbf{e} と \mathbf{f} は 3 次元ベクトルであり、いずれも未知とする。目的は、この μ を求めることである。

(18) 式は、 $\mathbf{G} - \mu\mathbf{I}$ がランク 1 であること、すなわち $\mathbf{G} = [\mathbf{a} \ \mathbf{b} \ \mathbf{c}]$, $\mathbf{I} = [\mathbf{i} \ \mathbf{j} \ \mathbf{k}]$ と書いた場合

$$\mathbf{a} - \mu\mathbf{i} \simeq \mathbf{b} - \mu\mathbf{j} \simeq \mathbf{c} - \mu\mathbf{k}$$

であることを意味する。よって、これら 3 つの列ベクトルからどの 2 つを選んでも、それらのベクトル積は 0 になる。例えば、最初の 2 つを取れば

$$(\mathbf{a} - \mu\mathbf{i}) \times (\mathbf{b} - \mu\mathbf{j}) = \mathbf{a} \times \mathbf{b} + \mu(\mathbf{j} \times \mathbf{a} - \mathbf{i} \times \mathbf{b}) + \mu^2 \mathbf{k} = \mathbf{0} \quad (19)$$

となる。(19) 式は、 μ に関して 3 つの拘束を与えるが、そのうち 2 つのみが独立である。そこで、(19) 式と \mathbf{i} および \mathbf{j} との内積をとって、

$$\mathbf{i}^T(\mathbf{a} \times \mathbf{b}) + \mu \mathbf{k}^T \mathbf{a} = 0, \quad \mathbf{j}^T(\mathbf{a} \times \mathbf{b}) + \mu \mathbf{k}^T \mathbf{b} = 0$$

なる μ に関して線形な 2 つの拘束を取り出す .

同様にして, 他の 2 組の列ベクトルの組み合わせから μ に関する 4 つの拘束が取り出せるので, 合計 6 つの拘束を得る . G にノイズが含まれなければ μ はこれらの共通根となるが, そうでない場合もこれらの最小自乗解として μ を求めることができる .

参考文献

- [1] O. D. Faugeras. *Three-Dimensional Computer Vision*. MIT Press, 1993.
- [2] A. Heyden. Projective Structure and Motion from Image Sequences Using Subspace Methods. In *Proc. 10th Scandinavian Conference on Image Analysis*, 1997.
- [3] A. Heyden and K. Åström. Euclidean Reconstruction from Image Sequences with Varying and Unknown Focal Length and Principal Point. In *Proc. Computer Vision and Pattern Recognition*, pp. 438–443, 1997.
- [4] F. Kahl. Critical Motions and Ambiguous Euclidean Reconstructions in Auto-Calibration. In *Proc. 7th International Conference on Computer Vision*, pp. 469–475, 1999.
- [5] S. J. Maybank and O. D. Faugeras. A theory of self-calibration of a moving camera. *International Journal of Computer Vision*, Vol. 8, No. 2, pp. 123–151, 1992.
- [6] M. Pollefeys, R. Koch, and L. Van Gool. Self-Calibration and Metric Reconstruction In spite of Varying and Unknown Intrinsic Camera Parameters. *International Journal of Computer Vision*, Vol. 32, No. 1, pp. 7–25, 1999.
- [7] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press, 1988.
- [8] P. Sturm. Critical Motion Sequences for Monocular Self-Calibration and Uncalibrated Euclidean Reconstruction. In *Proc. Computer Vision and Pattern Recognition*, pp. 1100–1105, 1997.
- [9] P. Sturm and B. Triggs. A Factorization Based Algorithm for Multi-Image Projective Structure and Motion. In *Proc. 4th European Conference on Computer Vision*, Vol. 2, pp. 709–720, 1996.
- [10] C. Tomasi and T. Kanade. Shape and Motion from Image Streams under Orthography: a Factorization Method. *International Journal of Computer Vision*, Vol. 9, No. 2, pp. 137–154, 1992.
- [11] B. Triggs. Autocalibration and the Absolute Quadric. In *Proc. Computer Vision and Pattern Recognition*, pp. 609–614, 1997.
- [12] R. Y. Tsai. A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses. *IEEE Journal of Robotics and Automation*, Vol. 3, No. 4, pp. 323–344, Aug 1987.
- [13] T. Ueshiba and F. Tomita. A Factorization Method for Projective and Euclidean Reconstruction from Multiple Perspective Views via Iterative Depth Estimation. In *Proc. 5th European Conference on Computer Vision*, Vol. 1, pp. 296–310, 1998.
- [14] Lihi Zelnik-Manor and Michal Irani. Multiview Constraints on Homographies. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 24, No. 2, pp. 214–223, Feb 2002.
- [15] Z. Zhang. A Flexible New Technique for Camera Calibration. Technical Report MSR-TR-98-71, Microsoft Research, Dec 1998.
- [16] Z. Zhang. A Flexible New Technique for Camera Calibration. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 11, pp. 1330–1334, Nov 2000.
- [17] 浅田尚紀. CVCV-WG 特別報告: コンピュータビジョンにおける技術 評論と将来展望 (XI) - カメラモデルとキャリブレーション手法 -. 情処研報, No. CVIM102–9, Nov 1996.
- [18] 出口光一郎, 岡谷貴之. 3次元環境理解のための多視点カメラのキャリブレーション. 情処研報, No. 2002–CVIM–131, pp. 1–8, January 2002.